Internet Bad Neighborhoods: the Spam Case

Giovane C. M. Moura, Ramin Sadre, and Aiko Pras
University of Twente
Centre for Telematics and Information Technology (CTIT)
Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS)
Design and Analysis of Communications Systems (DACS)
Enschede, The Netherlands
Email: {g.c.m.moura, r.sadre, a.pras}@utwente.nl

***Abstract*—A significant part of current attacks on the Internet comes from compromised hosts that, usually, take part in botnets. Even though bots themselves can be distributed all over the world, there is evidence that most of the malicious hosts are, in fact, concentrated in small fractions of the IP address space, on certain networks. Based on that, the Bad Neighborhood concept was introduced. The general idea of Bad Neighborhoods is to rate a subnetwork by the number of malicious hosts that have been observed in that subnetwork. Even though Bad Neighborhoods were successfully employed in mail filtering, the very concept was not investigated in further details. Therefore, in this work we provide a closer look on it, by proposing four definitions for spam-based Bad Neighborhoods that take into account the way spammers operate. We apply the definitions to real world data sets and show that they provide valuable insight into the behavior of spammers and the networks hosting them. Among our findings, we show that 10% of the Bad Neighborhoods are responsible for the majority of spam.**

I. Introduction

Over the recent years we have seen a shift on the way that malicious activities are performed on the Internet. While in the past most of the attacks originated from *single* compromised servers, a significant part of current attacks comes from *distributed* compromised machines, part of the so-called botnets [1]. Botnets can consist of millions of bots and have unprecedented networking and processing power. For example, by the end of 2010, the BredoLab botnet was estimated to have a spam capacity of 3.6 billion messages per day, by compromising more than 30 million hosts worldwide [2].

Even though bots can be distributed all over the world, there is evidence that most of malicious hosts are, in fact, concentrated in certain networks. For example, in 2006 Ramachandran *et al.* [3] have shown that the majority of spam was sent from a small fraction of the IP address space. Collins *et al.* [4], on the other hand, have defined the term "spatial uncleanliness" for clusters of compromised hosts. Also, DNS blacklists such as Spamhaus SBL [5] list entire netblocks[1] at different aggregation levels (e.g., /23, /24, /31), which also suggests the concentration of spamming hosts. To cope with this concentration, the *Bad Neighborhood* (BadHood) concept was introduced by Van Wanrooij and Pras [6]. The idea behind BadHoods is that the probability that a specific IP address behaves badly increases if neighbor IP addresses, *i.e.*, hosts within the same subnetwork, behave badly.

[1]We use the term netblock to specify network blocks at any aggregation level (e.g., /24, /32). Bad Neighborhoods, in turn, are netblocks that contain malicious hosts.

Although Bad Neighborhoods were introduced and employed for filtering out spam messages, the very concept itself was not investigated in further details. In this paper we focus on Bad Neighborhoods that send spam (*Spamming BadHoods* in the following), targeting the following research questions:

1) *What are the worst protected netblocks in the Internet?* – that is, netblocks containing a significant number of spamming hosts.
2) *What are the the most "spam-friendly" providers?*, *i.e*, providers that "turn a blind eye" to massive spammers in their networks [3].
3) *Do Spamming BadHoods with many spammers also send many spam messages?*
4) *How much data do we need to identify Spamming BadHoods?*

These research questions have led us to four distinct definitions for Spamming BadHoods. The first two take into account the two types of spammers observed on the Internet: Low-Volume Spammers (LVS) and High-Volume Spammers (HVS) [7]. The first type describes hosts "working under a central provision, each typically spamming with a low volume", while the second one consists of "dedicated spam sources, which are brute force spammers, each spamming in an enormous number every day". Since most of the LVS tend to be part of botnets [7], [8], concentrations of LVS reveal what are the "most infected networks" in the Internet – and, consequently, the worst protected ones. Therefore, this first definition addresses the first research question by identifying LVS BadHoods, *i.e.*, netblocks containing low-volume spammers.

To answer the second research question, we introduce a second definition, which covers Spamming BadHoods composed of HVS. This type of BadHood allows us to identify providers that ignore or tolerate dedicated spammers that spam at high volume in their networks and, therefore, can be considered as "spam-friendly". The third research question, on the other hand, leads to our next definition, which considers the total number of spam messages sent by each netblock. By comparing the number of spam messages and spammers per BadHood, we can determine if there is any correlation between

those two. Finally, the fourth research question leads to our last definition, in which we compare the BadHoods obtained from different data sources.

The rest of this paper is structured as follows. Section II details the four definitions of Spamming BadHoods. Next, Section III presents the datasets employed in our analysis. Following that, Section IV shows the experimental results. Related work is discussed in Section V. Finally, Section VI contains our conclusions and proposes future work.

II. Four definitions for Spamming Bad Neighborhoods

The Bad Neighborhood concept was first introduced in [6]. Even though the concept was successfully employed for mail filtering, it was not investigated in further details. In particular, Spamming BadHoods can be defined in several ways. In this section, we propose four definitions that allow us to gain more insight into the behavior of spammers and the networks hosting them. We begin with a brief discussion of the possible data sources to evaluate Spamming BadHoods in Section II-A. Then we present the four definitions of Spamming BadHoods from Sections II-B to II-E.

A. Possible Data Sources for BadHood Analysis

In order to identify and analyze Spamming BadHoods, we need to obtain the IP addresses of the spamming hosts. Several sources of data can be employed for that. Next we present an overview of those sources.

1) DNS Blacklists: One approach to identify IP addresses of spammers is to set up *spamtraps*, which are specialized honeypots to collect spam. By definition, every message that reaches a trap is considered spam, since it was unsolicited in the first place. Their source IP addresses can be used to build blacklists, usually in real time. The term "DNS Blacklist" comes from the fact that many blacklist maintainers allow queries to be made to their blacklists in a similar way DNS queries are performed. A comparison of blacklists can be found at [9] and [10]. Blacklists do not necessarily list the full IP address of every single spammer. In fact, some lists only provide aggregated information on whole subnetworks. Even though DNS blacklists list many IPs, they do not provide the information on *how many spam* messages a spammer has sent – they only tell that a certain IP has sent spam.

2) Mail Server Logs: Most of the spam is currently detected on mail servers, where incoming messages are processed and filtered. Mail filters, such as SpamAssassin [11], are configured to perform a series of checks on every e-mail message. These tests can include header and text analysis, Bayesian filtering, and even take input from DNS blacklists. Depending on the outcome of the tests, each message is classified as "spam" or "ham". Differently from DNS blacklists, it is possible to determine how many spam messages were sent by each IP address using mail server logs.

3) Mail Client Logs: Spam mails can also be identified by the mail client itself. This is usually the last resort against spam because it does not avoid the increased bandwidth usage caused by unsolicited mails. Similar to the mail filters used by mail servers, the mail client, such as Thunderbird [12], can perform a series of tests in order to classify the mails.

4) Network Flows: According to the IETF, a *network flow* is defined as a "set of IP packets passing an observation point in the network during a certain time interval that share the same properties" [13]. Typically, these properties include the source and destination port and address of the packets, as well as other IP header fields, like the protocol number. Flow probes monitor the packets in a network and export so-called *flow records*, which contain summarized information on the identified flows, such as the number of exchanged packets.

Flow monitoring can be used to detect spams [14], [15]. However, since flow records only provide an aggregated view to the network traffic, the validation of the detection results is much harder and has to be based on statistical arguments. Hence, we do not employ this data source in our analysis.

B. First Definition: LVS BadHoods

Low-Volume Spammers (LVS) are hosts that spam at a low volume to avoid being blacklisted. Typically, they are operated under a central provision, usually as part of a botnet [7]. The latter obviously requires that the host has to be firstly infected. Hence, a concentration of a high number of LVS in a subnetwork indicates that the particular subnetwork is poorly protected or managed, and that the responsible ISP migth neglect the malware propagation in their networks. The goal of this definition for Spamming BadHoods is to detect the worse protected (or infected) subnetworks by identifying the netblocks with many LVS. We refer to such BadHoods as LVS BadHoods.

The first step is to classify spammers according to the number of messages each of them has sent during the observation period. Note that this information is not provided by blacklists, so we have to rely on the other data sources. Since spammers can behave differently across different domains, we combine the data obtained from several observation points. After that, we need to establish a threshold θ that we apply to the number of sent messages in order to separate LVS from other spammers. We define:

$$\theta = d \times s \times m \tag{1}$$

In this equation, d is the length of the analyzed data trace in days, s is the number of different domains being monitored, and m is the maximum number of messages that a spammer can send to a single domain per day in order to be considered a LVS. As described in [7], [8], a LVS usually contacts a same mail server once or twice a day. Hosts spamming under this threshold are classified as LVS.

After classifying each host, we count the number of LVS per /24 netblock. For example, if the following set of /32 IP addresses were listed $\{1.1.1.5, 1.1.1.4, 1.1.1.64\}$, the block $1.1.1.0/24$ would have a count of three. The blocks are then ranked by their count (score). The maximum possible score per each /24 block is 254, since usually the addresses with .0 and .255 suffix are reserved for network identification and

broadcasting, respectively. Aggregating data on /24 level is preferable because this is the smallest prefix length that can be reverse delegated on the Internet [6].

C. Second Definition: HVS BadHoods

Complementary to LVS, High-Volume Spammers (HVS) are those hosts that spam in high volumes, *i.e.*, that send more messages during the observation period than specified by the threshold θ in Equation 1. As for LVS, blacklists cannot be used to identify HVS.

HVS are usually dedicated spamming hosts operated by professional spammers. Therefore, a high concentration of HVS in a particular subnetwork indicates that the ISP tolerates them. To identify HVS BadHoods, we follow the same approach employed in the second definition, that is, we count how many spam messages each spammer has sent during the observation period. Spammers above the threshold θ are considered HVS. After classifying each host, we count the number of HVS per /24 netblock and rank them according to that number. As in the first definition, the maximum score for a netblock is 254.

D. Third Definition: Spamming BadHood Firepower

The third definition for Spamming BadHoods focuses on evaluating the "firepower" of each netblock. Therefore, we identify the most spamming BadHoods on the Internet in terms of the number of sent spam messages – not on the number of spamming hosts. As for the two previous definitions, we have to rely only on mail server and clients logs as data sources, since DNS blacklists do not provide the needed information. The first step is to count how many spam messages each spammer has sent. Next, for each /24 netblock, we calculate the total number of spam messages sent by all the spammers located in the block. The final step is to rank the blocks according to that number.

E. Fourth Definition: All Spamming BadHoods

The goal of the last definition for Spamming BadHoods is to identify all spamming netblocks, independently of the spammers' behavior. Therefore, we need only the IP addresses of spammers, which allow us to include the information provided by DNS blacklists into our analysis. From each data source, we extract the IP addresses of the spammers. Then, all these IP are combined and duplicate entries are removed. Next, we count the number of spammers in each /24 block and rank the blocks according to the count.

III. Evaluated Datasets

In this section we present the data we have used in our experiments. We have obtained data from DNS blacklists, mail server logs, and mail client logs from various sources over a period of one week (April 19th-26th, 2010). Next we describe in more detail the collected data.

A. DNS Blacklists (DNSBL)

Table I shows the DNS blacklists we have obtained. The first one, Composite Blocking List (CBL) [16], maintains four large spamtrap infrastructures from where the source IP addresses of spammers are harvested. To give an idea of the size of their spamtraps, one of the four traps they maintain has received, on average, 2831.67 spams per second over a period of one year [17]. As shown in Table I, on April 21st more than 8 million unique IP addresses were listed on CBL.

Blacklist	Aggregation Level	# of entries (04/21/2010)
CBL	/32	8,334,895
PSBL	/32	2,445,270
UCEPROTECT-1	/32	3,350,417
UCEPROTECT-2	Various	30,143
UCEPROTECT-3	ASN	867
SBL	Various	10,535

TABLE I
DNS BLACKLISTS OBTAINED

Another blacklist we have obtained is the Passive Spam Block List (PSBL) [18]. PSBL is built using their own spamtraps, which capture around 500 thousands spam messages per day [19]. More than 2 million unique IP addresses were listed by PSBL on April 21st, 2010, as shown in Table I.

The next blacklists are from UCEPROTECT-Network [20]. They have three blacklists: The Level 1 blacklist lists only single IP addresses (/32) of spammers that have contacted their spamtraps. On April 21st, more than 3 million unique IP addresses were listed on the Level 1 blacklist. The Level 2 blacklist, on the other hand, is automatically generated based on Level 1. In that list, netblocks are entirely blacklisted according to a scoring procedure. The following entry is present on April 21st on the Level 2 blacklist: "*86.99.128.0/17 is UCEPROTECT-Level2 listed because 267 abusers are hosted by EMIRATES-INTERNET Emirates Internet/AS5384 there.*". Finally, the Level 3 blacklist lists all IP addresses from an autonomous system (except those whitelisted at ips.whitelisted.org) if "more than 100 IPs, but also a minimum of 0.2% of all IPs allocated to this ASN got Level 1 listed within the last 7 days".

Finally, the Spamhaus Block List (SBL) [5] lists IP addresses at different aggregation levels "from which Spamhaus does not recommend the acceptance of electronic mail" [5]. SBL contains single IP addresses as well as entire network blocks. As can be seen in Table I, more than 10 thousand entries are present on SBL on April 21st, 2010.

Since blacklists are usually built using a large number of honeypots, they have higher probability to be reached by many different spammers than a single mail server. For example, UT/EWI mail servers have been spammed by 71,754 different IP addresses on April 21st, 2010, while CBL spamtraps lists more than 8 million unique IP addresses. However, DNS blacklists lists only the IP addresses of spammers, while mail server logs list every single spam message – which allows to compute how many spam each spammer has sent.

B. Mail Servers Logs

Table II shows the mail servers from which we have obtained data. Provider A is a large hosting provider located in the Netherlands. Almost 7 million messages from more than 1.5 million different IP addresses were tagged as spam for the monitoring period (1 week). Next, we have obtained the log files from the mail server of the Electrical Engineering, Mathematics, and Computer Science Department at the University of Twente (UT/EWI)[2]. In total, more than 1.7 million messages were logged.

Domain	Country	# of Spam Msgs	# of distinct IPs
Provider A	The Netherlands	6,981,415	1,668,205
UT/EWI	The Netherlands	1,707,367	458,495
CAIS/RNP	Brazil	84,295	36,938
Provider B	France	1,160	975
Total	–	8,774,237	1,847,874

TABLE II
MAIL SERVERS LOG FILES ANALYZED

We have also obtained data from the mail servers of the Security Incident Response Team of the Brazilian Research Network (CAIS/RNP) [21]. More than 80 thousand spam messages were obtained for the monitoring period. Finally, we have obtained data from a small mail server hosted in France, denoted as Provider B, from which we got 1,160 spam messages from 975 distinct senders. In total, we have obtained more than 8.7 million messages from more than 1.8 million different IP addresses from mail servers.

C. Mail Client Logs

Finally, the last type of data collected was mail client spam logs. For this work we have obtained 1321 spam messages from 15 mail accounts from various countries. These messages, in turn, came from 763 different senders. Since this dataset is not as representative in comparison to the previous ones, we have not employed it in our analysis.

IV. Experimental results

In Section II we have introduce four definitions for Spamming BadHoods. In this section we apply the different definitions to the datasets presented in Section III and discuss the results.

A. LVS BadHoods

For this definition, we have combined mail server logs from four different domains ($s = 4$): Provider A, UT/EWI, CAIS/RNP, and Provider B. By combining those log files, we increase the chances of observing the same spammer on different mail servers which allows us to better classify it. The logs cover a period of seven days ($d = 7$). We choose $m = 2$ as the maximum number of mails sent by a LVS to a domain per day, as described in Section II-B. Hence, Equation 1 leads to a threshold θ of 56 spam messages, meaning that spamming hosts sending less or equal to 56 messages are classified as LVS while the others are HVS.

Table III shows the distribution of the number of spam messages per spamming hosts for the combined mail server logs. For a given number x of spam messages sent by a single spammer, the table gives the number of spammers (second column) that match it, followed by the total number of spams sent by those spammers (third column) and the average number of spams per spammer (fourth column). The numbers in parentheses give the percentages of the total number of spammers and spam, respectively, found in the dataset.

	# of IPs	# of Spam	Spam/IP
$x = 1$	867422 (46.94%)	867422 (9.8%)	1
$1 < x <= 10$	821472 (44.46%)	3189391 (36.35%)	3.88
$10 < x <= 56$	145648 (7.8%)	3053351 (34.8%)	20.96
$x > 56$	13081 (0.70%)	1662913 (18.95%)	127.12

TABLE III
DISTRIBUTION OF SPAM MESSAGES FROM MAIL SERVER LOGS (1 WEEK)

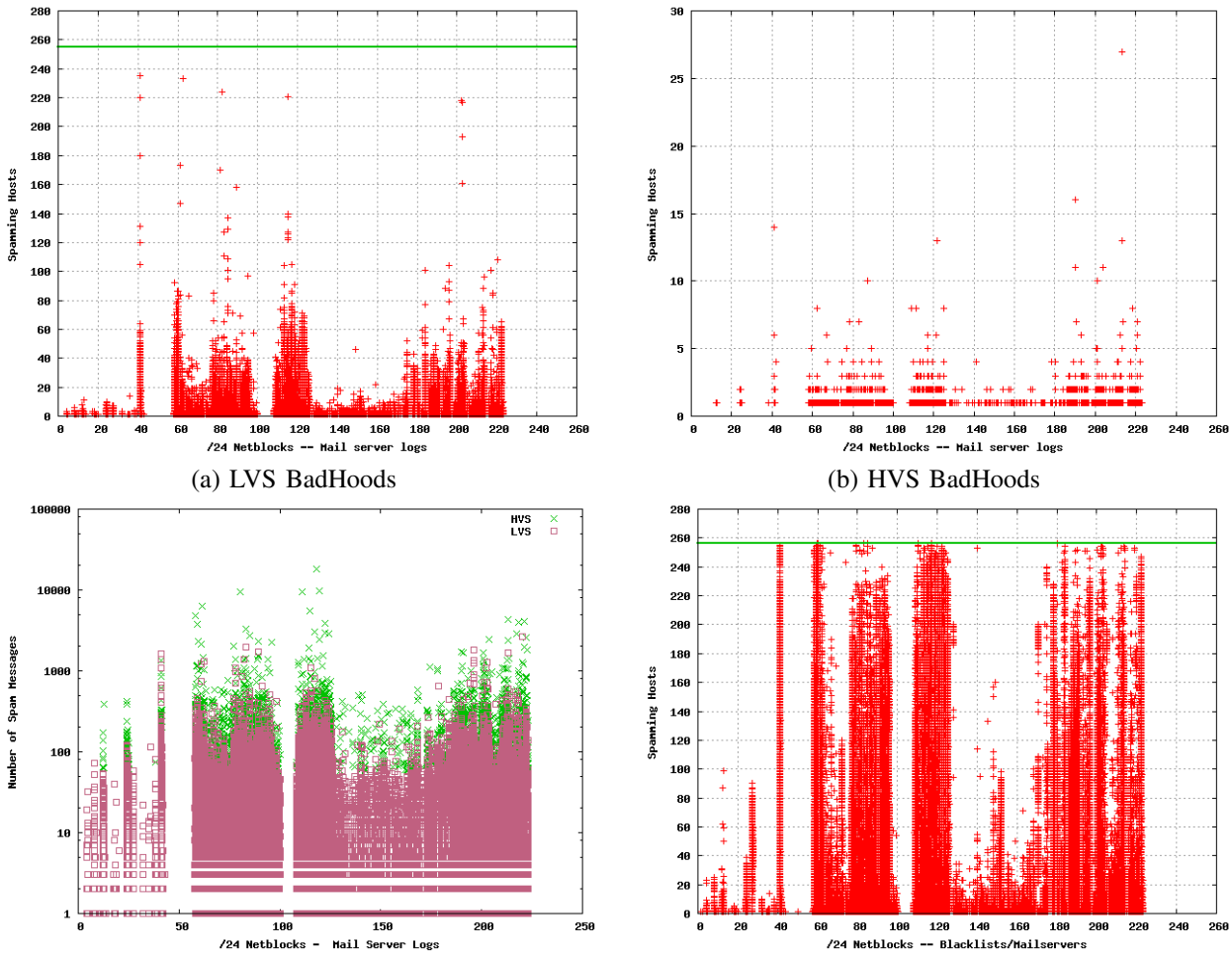
By employing the threshold θ of 56, one can observe that most of the spammers (99.3%) are classified as LVS (first to third rows in Table III) and that they are responsible for around 80% of all spam our mail servers have received. Since most of LVS are believed to be bots [7], [8], we can conclude from our results that probably most of the spam nowadays comes from botnets. In addition, we observe that nearly 50% of all spammers have only sent one message (first row), which confirms the tactic adopted by LVS to spam at very low volume to avoid being detected [7].

Figure 1(a) shows the distribution of LVS BadHoods over the IP address space. The x-axis gives the /8 prefix of the IP address; the y-axis gives the number of spamming LVS hosts per /24 block. Each point in the plot stands for one /24 block. The horizontal line shows the maximum possible number of hosts in a block, that is 254. We observe that there is a high concentration of spamming hosts on certain ranges, such as between 60-100, 110-125, and finally 180-200.

There are also IP ranges that we have not observed a single spammer. This includes the following ranges: 28-31, 44-49, 101-107. The reasons vary: some blocks were not allocated by IANA when we collected the data [22] (31,49,101-106). Others were legacy blocks (28-30, 45-47, 49), which were blocks usually assigned by the central Internet Registry (IR) prior to the Regional Internet Registries (RIRs). However, these blocks are managed by individual RIRs. Finally, the 46 block was allocated in September 2009, while 107 was allocated in February 2010.

For the blocks with the highest number of LVS spammers, we have identified the corresponding ISP. In the middle column of Table IV, we show the top 20 providers that manage the /24 most malicious LVS BadHoods. Analyzing this table, we could observe that some LVS BadHoods are

[2]UT/EWI has, in fact, a primary and a secondary mail server. We have combined the data from both as a single source. However, we should point out that spammers have targeted much more the secondary one. We think that spammers believe that secondary servers are not as secured as primary ones, which is not the case for UT/EWI servers.



(a) LVS BadHoods

(b) HVS BadHoods

(c) Spamming BadHoods Firepower

(d) All Spamming BadHoods

Fig. 1. Spamming BadHoods

made almost exclusively of spamming hosts. By evaluating only four different mail server domains, we were still able to find networks such as the case of Libyan Telecom having 235 spamming hosts in a single /24 netblock. Since most of LVS are believed to be bots, these results show how some ISPs neglect the propagation of bots and malicious activities that the hosts in their networks carry. This also confirms the facts that some DNS blacklists block entirely /24 netblocks, as SBL [5].

Another question one could raise is "if there is a correlation between the countries of the worse protected ISPs and their Internet legislation", or other variables, such as the usage of anti-virus in such countries. This, however, goes beyond the scope of this work. In our study, the top 20 malicious LVS BadHoods are located in Africa and Asia[3].

As explained in Section II-B, we can assume that the most malicious LVS BadHoods are also the worst protected and, consequently, the most infected ones. The presented results can be used by providers to raise awareness about the security of their networks and to improve it. In addition, LVS Bad Neighborhoods can also be employed to track and detect botnets [8].

B. HVS BadHoods

HVS BadHoods are determined in a similar way as the LVS BadHoods. We have employed the same datasets, period of time and threshold. As shown in Table III, only 13,081 (0.70%) IP addresses have been classified as HVS. The distribution of the HVS over the IP address space is visualized in Figure 1(b). Each point gives the number of HVS in one /24 block. We observe that most BadHoods host less than three HVS. Remarkably, some blocks contain up to 27 HVS – which is far less than for LVS cases, as shown in Table IV.

In the right column of Table IV, we show the top 20 "spam-friendly" providers. Differently from LVS BadHoods, we can find providers for HVS from Europe, Africa, Asia, Russia, and South America among the top 20. Even though the European Union has a directive that regulates spam [24], each member state is responsible for " *taking appropriate measures to ensure that [...] unsolicited communications for purposes of direct marketing [...] are not allowed either without the consent of*

[3]For a more detailed analysis on the geographical location of others types of attacks at city level, please refer to [23].

Rank	LVS BadHoods	HVS BadHoods
1	Libyan Telecom (235)	PTK Serbia (27)
2	Omantel Tech (233)	Digitel Venezuela (16)
3	Omantel Tech (224)	Maroc Telecom (13)
4	Digitel Philippines(221)	Smart Indonesia (13)
5	Libyan Telecom (220)	Vodafone Romenia (11)
6	Excelcomindo Philippines (218)	Smart Indonesia (11)
7	Smart Telecom Indonesia (217)	Digitel Venezuela (11)
8	Smart Telecom Indonesia (193)	Yahoo! Europe (10)
9	Libyan Telecom (180)	Telefonica Chile (10)
10	CAT Wireless Bangkok (173)	Maroc Telecom (8)
11	Maroc Telecom (170)	BSNLNET India (8)
12	Smart Telecom Indonesia (161)	Korea Telecom (8)
13	TATTELECOM Russia (158)	Chinanet (8)
14	CAT Wireless Bangkok (147)	Orange Romenia (8)
15	Digitel Philippines(140)	Orange Romenia (8)
16	Digitel Philippines(138)	OJSC MegaFon Russia (7)
17	OJSC MegaFon Russia (137)	Kazan Russia (7)
18	OJSC MegaFon Russia (137)	KORNET Korea (7)
19	Libyan Telecom (131)	KPN Netherlands (7)
20	OJSC MegaFon Russia (129)	CODETEL-Dominican Rep (7)

TABLE IV
PROVIDERS OF THE TOP 20 MOST MALICIOUS /24 NETWORKS (NUMBER OF HOSTS BETWEEN PARENTHESES)

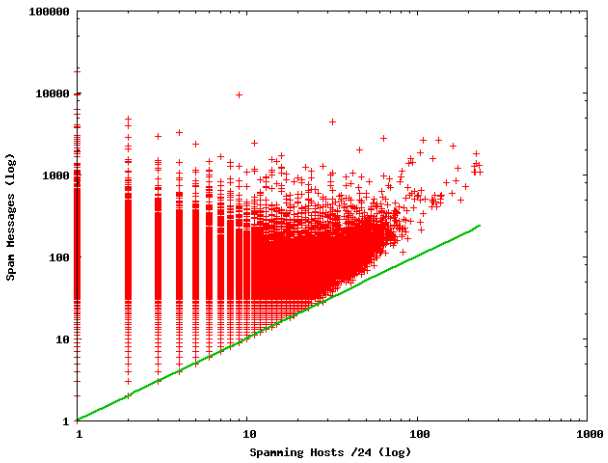


Fig. 2. Number of Spam Messages versus Number of Spamming Hosts per /24 block

the subscribers". Our results show that 5 of the top 20 HVS BadHoods are located within the EU borders, which raises doubts on the effectiveness of such directive.

Another interesting fact to observe is that Yahoo! Europe ranks number 8 in the Top HVS list. Checking manually the 10 IPs, we found out they are, in fact, mail servers from Yahoo! Mail located in the UK. This might be due account hijack, in which spammers hijack legitimate accounts to send spam [25], [26].

To conclude, HVS BadHoods show in which blocks HVS are located, thus allowing us to identify the most "spam-friendly" providers. The presented results can be used to raise awareness about those providers and, in some cases, alert to the number of hijacked mail accounts. This could be used by mail filters to appropriately rank mail from such netblocks and by ISPs to filter out outgoing email as well [27], [28].

C. Spamming BadHood Firepower

So far, we have analyzed BadHoods based on the number of spamming hosts per netblock. Therefore, in this section we evaluate neighborhoods in terms of their firepower, *i.e.*, their impact measured in number of spam messages they have sent. Again, we rely on the mail server logs to calculate the total number of spams sent per /24 netblock. The result is shown in Figure 1(c). Each point represents one /24 block. Note the logarithmic scale of the y-axis.

In Section I, we have raised the question whether the Spamming BadHoods with most spammers are also responsible for most of the spam. In Figure 2, we show the number of spams sent by a /24 block as function of the number of spamming hosts in that block. Each point represents one /24 block. The diagonal line in the plot visualizes the minimum number of spam messages that a netblock can send (which is equal to the number of spammers). We notice a large variation in the number of spam messages per netblock. The figure also shows that a higher number of spammers does not necessarily implies that a BadHood also sends more spams. Especially for blocks with up to ten spammers, there seems to even be a reverse relationship. Consequently, the Pearson correlation between the number of spamming hosts and the number of spam messages per netblock is quite weak with a coefficient of 0.32.

Our analysis also reveals how spam is distributed according to the Spamming BadHoods. Figure 3 presents the cumulative distribution function for the spam messages, where the BadHoods were ordered according to their firepower. As one can see, the majority of spam comes from a small fraction of all BadHoods. In fact, 10 % of the BadHoods were responsible for 54.87% of all spam. This suggests that, just by fighting a small subset of the malicious BadHoods, we should be able to block the majority of spam.

As a matter of fact, these results show the strength of the Bad Neighborhood concept. In Table III, we observed that 46.94% of spammers (/32 hosts) generate only 9.8% of the total amount of spam – which poses a major challenge for DNS blacklists-based spam detection. However, by employing

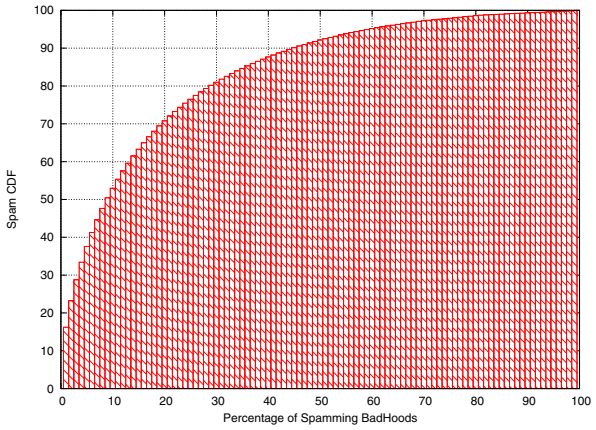


Fig. 3. Spam CDF

the Bad Neighborhood concept (/24 netblocks), we are able to invert this situation: we found that 10% of Spamming BadHoods were responsible for 54.87% of all spam.

Given the threshold of $\theta = 56$ spam messages, when ranking badhoods according to their firepower, many of them that have a small number of spammers must, therefore, contain HVS. In fact, among the top 10 worse Spamming BadHoods, six have only one spamming host, two have two spamming hosts, one has nine, and the last one has 32 spamming hosts.

This is also showed in Figure 1(c). We can see that the most evil Spamming BadHoods are, in terms of number of spam messages sent (or firepower), HVS BadHoods. However, most of BadHoods are classified as LVS. Even though the HVS BadHoods are minority, what we can learn from these results is to not underestimate the HVS BadHoods firepower and the damage that they can incur. On the other hand, the average firepower of the individual LVS BadHoods is lower. LVS BadHoods become powerful through their sheer number.

D. All Spamming BadHoods

As explained in Section II-E, the goal of the fourth definition for Spamming BadHoods is to identify all spamming netblocks, independently of the spammers' behavior. For our analysis, we used the data from the /32 blacklists (CBL, PSBL, UCEPROTECT-1) and the mail server logs (Provider A, UT, CAIS/RNP, Provider B) covering the period of April 19th to April 26th, 2010. This resulted in a list of more than 124 million entries with 15 million unique IP addresses. We aggregated the data by counting the number of spammers for each /24 block. By doing this, we found 1,205,888 /24 netblocks with at least one spammer. Figure 1(d) shows the distribution of the spammers over the IP addressing space. Each point gives the number of spammers of one /24 block. The x-axis specifies the /8 prefix of the blocks.

Our main motivation for this definition is the question, how much data is actually needed to identify all Spamming BadHood. Comparing Figure 1(d) with Figure 1(a), which has been generated only using mail server logs, we observe a similarity between the results. In both figures, we can identify the same regions with high, respectively low, activity. In total, mail server logs have allowed us to identify 571,389 Spamming BadHoods, while DNS blacklists combined together with mail server logs allowed to detect 1,205,932 Spamming BadHoods. The difference of 634,543 between the two shows how much extra BadHoods have been identified by using the additional information from DNS blacklists.

However, we should put this into perspective: the blacklists we have used in our analysis have provided more than 115 million entries, while the mail servers have provided 8.7 million. But, by using the blacklists, we were able to identify only 2.11 times more BadHoods. We can conclude that even though blacklist provide much more data, mail server logs perform quite well when finding Spamming BadHoods.

V. Related Work

Several research works have suggested that malicious hosts are concentrated on some subnetworks on the Internet. The network level behavior of spammers was analyzed in [3]. The authors have collected spam from a spam sinkhole for more than one year. They have shown that most of spam comes from a few concentrated part of IP address space. In our work, however, we have obtained spam from mail servers from production networks – which have been running legitimate mail servers for years – and DNS blacklists. DNS blacklists, such as [5], [16], [18], list malicious IP addresses at different aggregation levels, suggesting concentration on some sub-networks. In [4], the authors have defined the concept of uncleanliness, that "works as an indicator for how likely the network is to contain compromised hosts".

In another work [7], the authors have set up an open relay for e-mail and proposed a classifcation for spammers we have employed in this work: low-volume spammers (LVS) and high-volume spammer (HVS). We have extended this concept to BadHoods and employed it in our analysis. In [29], the authors have conducted an experiment by becoming part of a spamming botnet. They were able to observe a spam campaign over a period of week, in which 400 million spam messages were sent.

The Bad Neighborhood concept was introduced in [6]. In that study, the authors collected data from several DNS blacklists and counted by the number of spamming hosts per block that were identified those blacklists. The resulting count was then transformed into a score for the /24 netblock. Together with other data, they employed this score to determine whether an e-mail would be spam or not based on its sender's IP. If the message originated from a "bad neighborhood", *i.e.*, from a /24 netblock with a high score, it was rated as spam, even if the particular sender IP was never observed as spammer before. Our work goes beyond the previous one by defining and analyzing four types of Spamming BadHoods. The results provided in our work can be used to further develop current detection algorithms.

VI. Conclusions and Future Work

In this paper we have investigated Internet Bad Neighborhoods, focusing on Spamming Bad Neighborhoods. We have raised four research questions that led us to four definitions for Spamming BadHoods.

The first research question was "*What are the worst protected netblocks*". We have defined and evaluated LVS BadHoods to answer this question. The results have shown some ISPs seem to completely neglect the malware propagation in their networks. We can use these results to raise awareness on the most infected networks and to improve security levels on those networks. Such BadHoods can also be employed to characterize botnets.

The second research question addressed was "*what are the most spam-friendly providers*". We have defined HVS BadHoods to answer this question, and identified such friendly providers. The presented results can be used to raise awareness about those ISPs and to question the effectiveness of Spam legislation, such as the EU directive 2002/58 [24] The HVS

information could be used by mail filters to block, or at least appropriately rank, mails from such BadHoods.

The third research question was whether "*Spamming BadHoods with many spammers also send many spam messages?*". Our analysis revealed that this is not the case. In fact, the top 10 Spamming BadHoods had no more than 32 spamming hosts (and 6 of them had only one, including the most active BadHood). In addition, we have shown that most of spam comes from a fraction of all BadHoods. The lesson we can learn is that we should not underestimate HVS BadHood firepower. And that the list of HVS providers should be used to raise awareness and improve security of such providers.

Finally, the last question addressed in this paper was "*how much data do we need to find Spamming BadHoods*". We have shown that DNS blacklist help to obtain twice as much BadHoods than when only relying on our mail server logs. However, they have listed 13 times more IPs. We can conclude that even though blacklists provide more data, mail server logs perform quite well when finding Spamming BadHoods.

For more than 15 years, the Internet community has been fighting against spam, and the problem still far from being solved. In this work, we have provided an insight on Spamming BadHoods, taking into account spammers behavior and firepower. The definitions and results presented can be used to refine current solutions to fight spam.

As future work, we intend to extend our analysis to different types of attacks, such as SSH-scanning Bad Neighboorhoods and Distributed Denial of Service (DDoS) BadHoods. The idea is to observe if they exhibit the same behavior as Spamming BadHoods. Also, we intend to do the same analysis over extended periods of time in order to observe if there are any temporal patterns on the way BadHoods operate.

Acknowledgments: This research work has been partially supported by the EU UniverSelf Collaborative Project (#257513). The authors would like to thank Anna Sperotto, Casper Eyckelhof, Jáder Moura, Jéferson Nobre, Jürgen Schönwälder, Jürgen Rochol, Lisandro Granville, Luciano Gaspary, Marc Berenschot, Marijn Jongerden, Nikolay Melnikov, Olivier Festor, Pieter-Tjerk de Boer, Stephan Roolvink, Tiago Fioreze, Tobias Bandh, Ward van Wanrooij, and Wouter de Vries for their spam and support for this research. Special thanks to Frederico Costa and Liliana Solha from CAIS/RNP.

REFERENCES

[1] E. Cooke, F. Jahanian, and D. McPherson, "The zombie roundup: understanding, detecting, and disrupting botnets," in *Proceedings of the Steps to Reducing Unwanted Traffic on the Internet Workshop*. Berkeley, CA, USA: USENIX Association, 2005.

[2] InfoSecurity.com, "BredoLab downed botnet linked with Spamit.com," November 2010. [Online]. Available: http://www.infosecurity-magazine.com/view/13620/bredolab-downed-botnet-linked-with-spamitcom

[3] A. Ramachandran and N. Feamster, "Understanding the Network-level Behavior of Spammers," in *Proceedings of the 2006 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, ser. SIGCOMM '06. New York, NY, USA: ACM, 2006, pp. 291–302.

[4] M. P. Collins, T. J. Shimeall, S. Faber, J. Janies, R. Weaver, M. De Shon, and J. Kadane, "Using Uncleanliness to Predict Future Botnet Addresses," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, ser. IMC '07. New York, NY, USA: ACM, 2007, pp. 93–104.

[5] The Spamhaus Block List, 2011. [Online]. Available: http://www.spamhaus.org/sbl/

[6] W. van Wanrooij and A. Pras, "Filtering Spam from Bad Neighborhoods," *International Journal of Network Management*, vol. 20, no. 6, pp. 433–444, November 2010.

[7] A. Pathak, Y. C. Hu, and Z. M. Mao, "Peeking into Spammer Behavior from a Unique Vantage Point," in *Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*. Berkeley, CA, USA: USENIX Association, 2008.

[8] L. Zhuang, J. Dunagan, D. R. Simon, H. J. Wang, and J. D. Tygar, "Characterizing Botnets from Email Spam Records," in *Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*. Berkeley, CA, USA: USENIX Association, 2008.

[9] J. Makey, "Blacklists Compared," 2011. [Online]. Available: http://www.sdsc.edu/~jeff/spam/Blacklists_Compared.html

[10] Blacklist Statistics Center, 2011. [Online]. Available: http://stats.dnsbl.com/

[11] SpamAssassin, "The SpamAssassin Project," 2011. [Online]. Available: http://spamassassin.apache.org/

[12] Thunderbird, "Junk Mail Controls," 2011. [Online]. Available: http://kb.mozillazine.org/Junk_Mail_Controls

[13] J. Quittek, T. Zseby, B. Claise, and S. Zander, "Requirements for IP Flow Information Export (IPFIX)," RFC 3917 (Informational), Internet Engineering Task Force, Oct. 2004.

[14] A. Sperotto, G. Vliek, R. Sadre, and A. Pras, "Detecting spam at the network level," in *Proceedings of the 15th Open European Summer School and IFIP TC6.6 Workshop on The Internet of the Future*, ser. EUNICE '09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 208–216.

[15] A. Sperotto, G. Schaffrath, R. Sadre, C. Morariu, A. Pras, and B. Stiller, "An Overview of IP Flow-Based Intrusion Detection," *Communications Surveys Tutorials, IEEE*, vol. 12, no. 3, pp. 343 –356, 2010.

[16] The CBL, 2011. [Online]. Available: http://cbl.abuseat.org/

[17] CBL, "Spam Trap Flow Statistics," 2011. [Online]. Available: http://cbl.abuseat.org/totalflow.html

[18] Passive Spam Block List, 2011. [Online]. Available: http://psbl.surriel.com/

[19] PSBL FAQ, 2011. [Online]. Available: http://psbl.surriel.com/faq/

[20] UCEPROTECT-Network - Germany's first Spam protection database, 2011. [Online]. Available: http://www.uceprotect.net/en/

[21] CAIS, "Security Incident Response Team (In Portuguese: *Centro de Atendimento a Incidentes de Segurança*," 2011. [Online]. Available: http://www.rnp.br/en/cais/

[22] IANA IPv4 Address Space Registry, 2011. [Online]. Available: http://www.iana.org/assignments/ipv4-address-space/ipv4-address-space.txt

[23] M. van Polen, G. C. M. Moura, and A. Pras, "Finding and Analyzing Evil Cities on the Internet," in *Proceedings of the 5th International Conference on Autonomous Infrastructure, Management and Security (AIMS 2011)*, vol. 6734. Nancy, France: Springer Verlag, 2011.

[24] EU, "Directive on Privacy and Electronic Communications (2002/58)," July 2002. [Online]. Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002L0058:EN:NOT

[25] Comtouch, "July 2011 - Internet Threats Trend Report," July 2011. [Online]. Available: http://www.commtouch.com/download/2085

[26] Dick Craddock, "Hey! My friend's account was hacked!" 2011. [Online]. Available: http://windowsteamblog.com/windows_live/b/windowslive/archive/2011/07/14/hey-my-friend-s-account-was-hacked.aspx

[27] W. W. de Vries, G. C. M Moura, and A. Pras, "Fighting Spam on the Sender Side: A Lightweight Approach," in *Proceeings of 16th EUNICE/IFIP WG 6.6 Workshop (EUNICE 2010)*, ser. Lecture Notes in Computer Science, F. A. Aagesen and S. J. Knapskog, Eds., vol. 6164/2010. Berlin: Springer Verlag, 2010, pp. 188–197.

[28] J. François, G. C. M. Moura, and A. Pras, "Cleaning Your House First: Shifting the Paradigm on How to Secure Networks," in *5th International Conference on Autonomous Infrastructure, Management and Security (AIMS 2011), Nancy, France*, I. Chrisment, A. Couch, R. Badonnel, and M. Waldburger, Eds., vol. 6734. Berlin: Springer Verlag, June 2011.

[29] C. Kreibich, C. Kanich, K. Levchenko, B. Enright, G. M. Voelker, V. Paxson, and S. Savage, "On the Spam Campaign Trail," in *Proceedings of the 1st Usenix Workshop on Large-Scale Exploits and Emergent Threats*. Berkeley, CA, USA: USENIX Association, 2008.